*Environment and Planning A* 2007, volume 39, pages 630 – 646

DOI:10.1068/a37441

# Urban spatial change and excess commuting

**Kang-Rae Ma**
Centre for Transport and Society, Faculty of the Built Environment, University of the West of England, Bristol BS16 1QY, England; e-mail: kang.ma@uwe.ac.uk

**David Banister**
The Bartlett School of Planning, University College London, London WC1H 0QB, England; e-mail: d.banister@ucl.ac.uk
Received 22 December 2004; in revised form 27 May 2005

**Abstract. In this paper we revisit the excess-commuting technique and its links with urban form. The uncertainty in measurement is highlighted, as are the problems relating to changes in excess commuting over time. The measure of the theoretical maximum commute is proposed and added to the traditional excess-commuting measure so that the use of both the minimum and maximum levels can capture the concept of commuting potential. This measure is what we call the 'extended excess-commuting measure'. These concepts are tested through the use of a simulation exercise. As well as arguing for the inclusion of socioeconomic variables in analysis, we demonstrate that decentralisation in urban spatial structure can lead to either an increase or a decrease in average commuting distance. Some of the inconsistencies in the use of excess commuting can be reduced through the use of actual commutes together with the commuting range, as these factors in combination lead to a clearer understanding of commuting efficiency.**

## 1 Introduction

Excess commuting can be calculated from the difference between the average actual commute and the minimum (optimal) average commute in the standard linear-programming transport-problem procedure. It has been extensively studied over the past two and a half decades, as it provides useful policy implications and insights into both the urban-travel efficiency levels and the potential commuting-travel savings that could be obtained given the existing jobs and housing-location distribution.

The excess-commuting values obtained for different cities have been compared in many previous studies (for example, Cropper and Gordon, 1991; Frost et al, 1998; Giuliano and Small, 1993; Hamilton, 1982; 1989; Horner, 2002; Kim 1995; Merriman et al, 1995; Scott et al, 1997; White, 1988). In this context, it is surprising that this method has been given little attention with respect to an application tool for benchmarking commuting efficiency of a particular city over different time periods. The one exception here is Frost et al (1998) where the analysis is applied at two points in time. But more research and a deeper understanding are still needed for the analysis of excess-commuting change over time. In this paper we try to provide such an understanding in terms of different urban forms and its links with the excess commuting so that the subtleties of the ways in which excess commuting changes can be assessed. Our first aim is to explain the uncertainty in the measurement of excess commuting when it is applied to different time periods, and the second aim is to investigate how this measure could possibly be utilised more effectively through an incorporation of the theoretical maximum-commute concept into the conventional excess-commuting measure.

In the first section, we discuss why the geographical distribution of jobs and residential locations is important in the study of workers' travel behaviour. The complex nature of the urban form and travel patterns is explained within the framework of

Brotchie's urban-triangle model. We then widen the debate by linking the concept of Brotchie's triangle to the excess-commuting method. The concept of urban-commuting potential is introduced to understand how the range of urban-commuting potential varies according to the urban form. Excess-commuting changes are demonstrated through eleven scenarios that outline the different combinations of changing minimum-commute distances and actual-commute distances. The implications of each scenario are explained with respect to the previous literature. Six simulations are presented to investigate the relationship between land-use patterns and spatial interaction, and they are then used to explain a variety of hybrid urban forms. In the final section we summarise the main findings and discuss future research directions.

## 2 Basic concepts

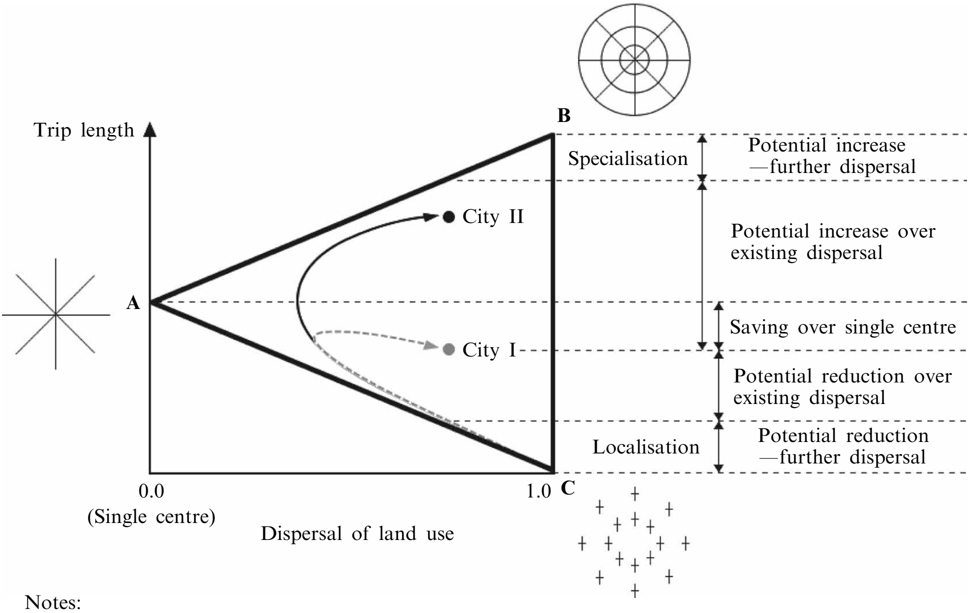
### 2.1 Urban spatial change and travel

There is a complexity over the links between the components of urban spatial structure. Bourne (1982) suggests that these links can be analysed through an understanding of urban form and spatial interaction, as they are organised by a set of rules, such as those describing the land market and the planning process. Yet all of these factors are highly intercorrelated, and so it is difficult to unravel the underlying dimensions. The theoretical simplicity of such concepts is also compounded by a reality that demonstrates the vast differences that exist between individual cities. It is probable that other specific factors such as sociodemographic differences also are correlated to the complex patterns of urban spatial structure. This may be part of the reason why the relationship between urban form and spatial interaction has not been clear despite an extensive body of literature on the association between urban form and travel (see Crane, 2000). No detailed review of studies on the relationship between the urban form and travel is given here, as it is not our main focus in this paper (see Banister, 2005a).

In understanding the relationship between urban form and spatial interaction, the dynamics of employment and population in most metropolitan cities have raised the question of how these are associated with travel behaviour. Gordon and Wong (1985) argued that further transformation to a polycentric city creates shorter journey-to-work trips. In addition, in a recent US study, Crane and Chatman (2003) suggested that the marginal effect of the decentralisation of jobs seems to be associated with shorter distance commutes and other factors appear to be important in explaining the generation of longer trips to work. If this is the case, it is considered that the emergence of local centres has several merits in terms of urban travel as they may help alleviate traffic congestion through a reduction in commuting journey distances (O'Sullivan, 1999).

However, this assertion may only be supported when firms and residences are locating relatively close to each other in the process of decentralisation. The simple triangle model by Brotchie (1984) provides a useful framework for the analysis of the process of decentralisation and travel. Even though his original intention of this model was to explain the impact of technological change on urban activities and interactions, his work provides very useful insights into the relationship between urban spatial change and urban travel. Figure 1 shows the relationship between the degree of land-use dispersal and trip length.

In figure 1, the vertical axis represents trip length. The horizontal axis shows the level of dispersal, which ranges from 0 for concentration in the centre to 1 for evenly distributed land uses. But the horizontal axis can be interpreted as a measure of city size if we expect that the transformation to a polycentric city may occur as the size of




Notes:
(1) The figure was slightly modified by the authors of this paper to address clearly the potential increase (or reduction) over existing dispersal.
(2) Only physical movements of workers within a city are considered in the figure.
(3) If the horizontal axis is interpreted as a city size, as in Newman and Kenworthy (1992, page 358), point A starts from origin (0.0) in the graph, thus the shape of the graph is ⦣. In this case, the triangular shape can be described by ◢

**Figure 1.** The triangle showing the relationship between land-use dispersal and journey-to-work trip length (source: Brotchie et al, 1996, page 88).

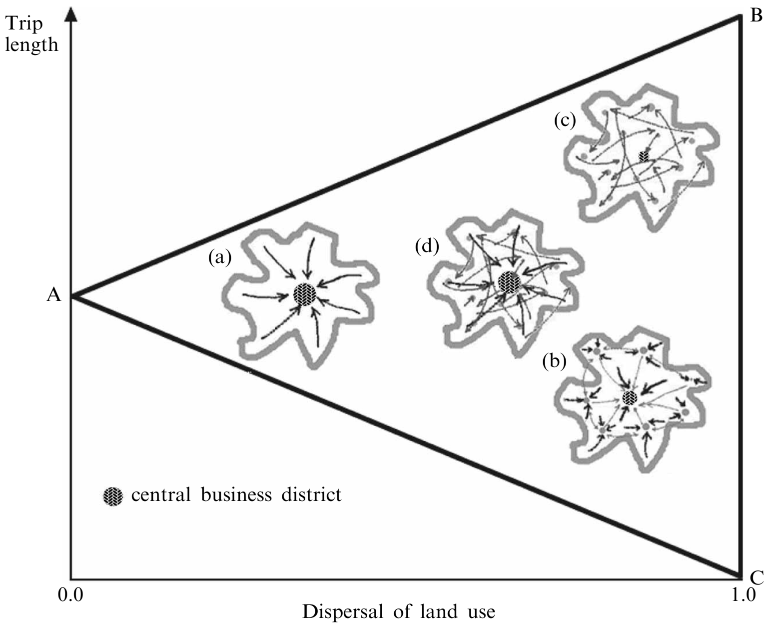
the city increases (see Newman and Kenworthy, 1992 for the relationship between city size and travel distance).

Point A in figure 1 represents the monocentric city in which all jobs are at the centre. Both points B and C represent the city form where there is a complete dispersal of jobs and population (all jobs are equally dispersed, like the population). Point B represents complete dispersal of activities in which people choose their locations without regard for physical distance between jobs and housing and a majority of workers go to distant work places by bypassing nearby jobs. It may be interpreted that the line AB represents the situation in which there are large numbers of cross-commutes, resulting in very long journey-to-work trip distances in the existing land-use patterns. In contrast, point C represents preindustrial societies, where transport or communications networks are minimal. Therefore, the line AC shows the situation in which most workers choose nearby jobs (it should be noted that Brotchie's interpretations of the lines AB and AC are different from ours where only physical movements are considered. In the original triangle model, the maximum average trip length or interaction can be achieved when the unit trip cost is zero, thus the line AB represents this limit. In contrast, the minimum average trip length is concerned with the line AC in which there is extremely expensive or difficult travel per unit distance. However, Brotchie (1984) noted that a movement of the transport system towards the line AC could represent the increase in efficient movements of goods and people).

As a simple example, the comparison between city I and city II reveals that city I has been developed in a more efficient way in terms of journey-to-work lengths compared with city II, even though both cities have the same degree of dispersal. As far as city I is concerned, the dispersed city structure has an advantage over the monocentric city structure, as the trip lengths are shorter, whereas in city II the opposite is true. The further dispersal of employment creates the possibility of reducing workers' journey-to-work trip distance even more, and city I has the potential for reducing existing dispersal levels. On the other hand, the dispersal of employment can create the possibility of an increase in journey-to-work trip length if there is an increase in cross commuting. In this case (city II), a dispersed city may have longer commuting journey lengths as compared with a single-centred city.

The urban patterns are the outcomes of decisions by various actors such as households, firms, and public-sector agencies. The patterns of the city may be different according to the city's economic, political, or geographical situations. They might lead to the different use of travel modes such as car, public transit, walking, and cycling. In addition, commute patterns may be different according to a number of sociodemographic factors such as income, race, sex, and workers' preferences. For this reason, one cannot easily categorise the relationship between urban structure and travel patterns. There are many hybrids in the real world. Commuting distances could be different even though the cities are of a similar size in terms of population and employment. Figure 2 illustrates the complexity of the urban spatial structures and travel patterns.

In figure 2, both city (a) and city (d) can be described as urban areas that have a high degree of employment in the city centre. In city (a) (figure 2), under the monocentric urban pattern with completely centralised employment, shortening commuting journey length does not seem to be an easy option as all workers need to commute to the central city in which all jobs are located. City (d) shows the monocentric model with simultaneous radial and random movement and this pattern of urban structure is characteristic of many large cities.




Notes:
In the figure, four different trip patterns within a metropolitan area are taken from Bertaud (2002). Bertaud described city (a) as the monocentric model, city (b) as the polycentric model (the urban-village version), city (c) as the polycentric model (the random-movement version), and city (d) as the monocentric model (simultaneous radial and random movement), respectively.

**Figure 2.** A number of hybrids of urban spatial structure and travel patterns.

Both city (b) and city (c) show polycentric urban structures. However, there is a huge difference in average trip length between city (b) and city (c) despite the same degree of dispersal of land use. They are the clear examples that the urban flow is not determined solely by urban form. Two urban areas with the completely same distribution of jobs and housing may have very different journey-to-work travel patterns and lengths, because of commuters' other characteristics, such as social and ethnic status and different preferences (Anderson et al, 1996). There is a large difference in the potential for trip-length reduction over the existing land-use patterns, with city (c) having the largest potential reduction, followed by city (d) (figure 2).

Brotchie et al (1996) made a comparison of the major US, UK, and Australian cities and found that many metropolitan cities brought about further trip shortening over a single-centred city. These cities included the three British examples (London, Manchester, and Birmingham) and the five Australian case-study cities (Brisbane, Sydney, Melbourne, Adelaide, and Perth). On the other hand, they also showed that the average trip length varies considerably between US cities. For example, Baltimore, San Antonio, and Rochester have longer journey lengths, whereas Boston and Los Angeles have shorter journey lengths. Whether the selected metropolitan cities used in Brotchie et al (1996) conform to the trip patterns described in figure 2 needs further investigation and discussion (for example, the comparison of different cities should be made in relation to their size). Through their study Brotchie et al (1996) have made several important contributions to the existing literature by giving a useful insight into the relationship between land-use dispersal and trip length. It provides the basis for the discussion in the remaining sections in this paper.

The implications of the discussion in this section may be summarised as follows: (1) in some cases, it may be possible to bring about efficient commuting patterns without any changes in urban form; (2) dispersed or polycentric urban structure provides further potential for reducing commuting journey lengths; (3) Conversely, the transformation from a monocentric to a dispersed urban structure provides further potential for increasing commuting journey lengths.

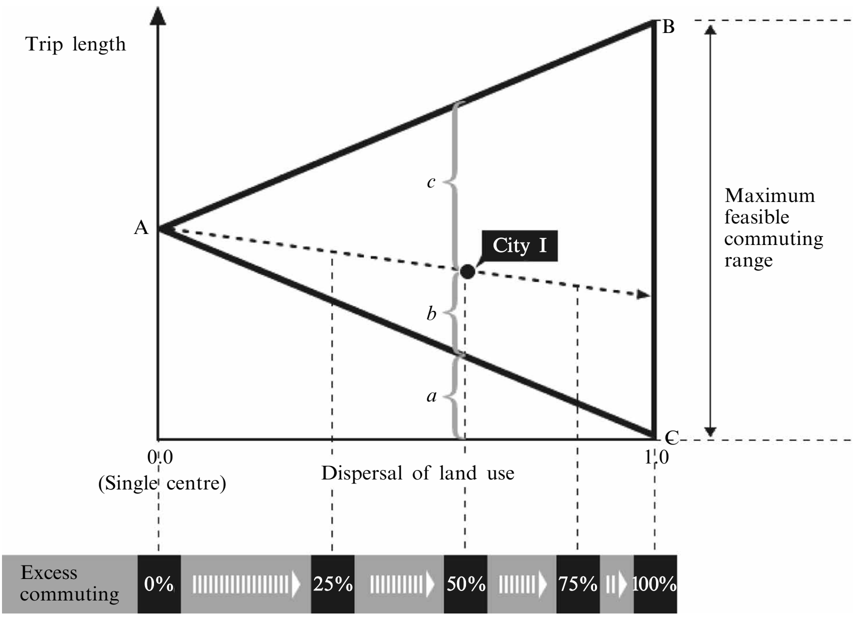
### 2.2 Urban spatial change and excess commuting

Brotchie's triangle construct shows conceptually the amount of potential reduction over existing dispersal (figure 1), and it can be conveniently linked to the excess-commuting technique. The concept of potential trip-length reduction over existing land-use dispersal is similar to that of excess commuting, and this can be calculated by reassigning workers to houses given the actual fixed locations of houses and jobs in the urban area, so as to minimise total commuting in the standard linear-programming transport-optimisation-problem procedure. The difference between the average actual and the theoretical minimum distances is excess commuting, and it can be reduced by trading jobs and housing, without restructuring the current land-use patterns (see the appendix).

The concept of potential trip increase over existing land-use dispersal cannot be described by the standard excess-commuting method as this is concerned only with the average actual commute and the average minimum commute. The introduction of the new concept, the theoretical maximum commute, can provide additional information on the potential trip-length increase, and this allows the feasible commuting range to be calculated (see the appendix). The feasible commuting range can be obtained by calculating the difference between the theoretical minimum and maximum commutes.

Furthermore, when the concept of the theoretical maximum commute is incorporated and used over multiple time periods, this can show the relationship between land-use dispersal and trip length (this is similar to Brotchie's triangle). The average maximum commute can be obtained by maximising total commuting in the linear-programming transport problem. Such an extended excess-commuting approach has been developed by Horner (2002) and Ma (2002), and it can provide the information about both the feasible best and worst scenarios in a given city form (note that the extended commuting measure is based on two different optimisation problems to estimate a minimum commute and a maximum commute). The extended excess-commuting method (see Ma, 2004) can demonstrate both the best scenario in which most workers choose nearby jobs in a given city's form and the worst scenario in which there are extreme cross-commutes. Horner (2002) named the difference between best and worst scenarios the 'commuting potential' or the 'carrying capacity', as the average actual commuting value must fall between these limits.

As can be seen in figure 3, the more dispersed city structure has the flexibility to either reduce or increase the average commuting distances by moving or swapping jobs and residential locations. The difference between the theoretical minimum commute and the theoretical maximum commute ($b + c$ in figure 3) becomes larger as the city becomes more decentralised. Because of a larger commuting potential range, even though average commuting is equal, the more dispersed cities are likely to produce more excess commuting. As an example, if a certain city has the urban development trajectory of city I in figure 3, this city has the advantage over the single-centred city in terms of average commuting distance. In the case of city I, however, the amount of excess commuting shows a gradual increase despite a decrease in average commuting distance. This is the reason why the standard measure of excess commuting is not helpful in investigating changes of urban travel efficiencies between two different time points, even though excess commuting may offer information about commuting efficiencies at one point in time.




Notes:
(1) In the figure above, the proportion of excess commuting can be calculated by $[b/(a+b)]$.
(2) The urban commuting potential can be expressed by $(a+b+c)$.
(3) The illustrated location of city I in the trajectory is where $a = b$.

**Figure 3.** Urban spatial change and excess commuting.

Urban commuting efficiencies can be discussed with respect to urban commuting potential which is determined by the geographical distribution of jobs and housing. If urban form is not taken into account in the excess-commuting measure, the results may be seriously misleading in comparative studies both between different cities and between different time periods.

## 3 The measurement of urban spatial change and excess commuting

Many cities have changed from a monocentric to a decentralised (for example, a polycentric or a dispersed) structure over the past few decades. Most studies carried out both on North American metropolitan cities including Chicago, Dallas-Fort Worth, Los Angeles, San Francisco, and Montréal and on other major Asian and European cities have demonstrated that there has been a further transition to a polycentric city (Baumont et al, 2003). Many researchers have questioned the validity of the monocentric urban model because its predictability has become unstable when applied to recent city development. This growing interest in a metropolitan area with multiple subcentres, and the theoretical and empirical models of polycentric structure are well discussed in Anas et al (1998).

This transformation has very important transport implications in that commuting behaviour is highly correlated to the physical distribution of employment and population. The transformation to a polycentric city, on the one hand, may shorten journey-to-work trip length by locating firms closer to the region where many workers are living. On the other hand, it may create a lot of cross commuting as a result of the difference in characteristics between residents and workers within an area, therefore lengthening commuting trip distance. This transformation from a monocentric to a polycentric structure may be associated with the suburbanisation process. However, a city may be experiencing a reurbanisation process in which a decline of population and jobs in the periphery takes place, with the core experiencing a corresponding increase in population and/or jobs (for an extended discussion of the conceptual urbanisation cycle see Klaassen and Paelinck, 1981).

The important point here is that all these dynamic situations could affect the estimation of excess commuting. In this section, all eleven possible scenarios of the effects of changing urban form in relation to the average minimum values on the excess commuting measure are summarised, and then the implications of each scenario are explained, using a British case study from Frost et al (1998). In addition, six simulations are conducted to investigate further the relationship between changing urban form and urban commuting capacity.

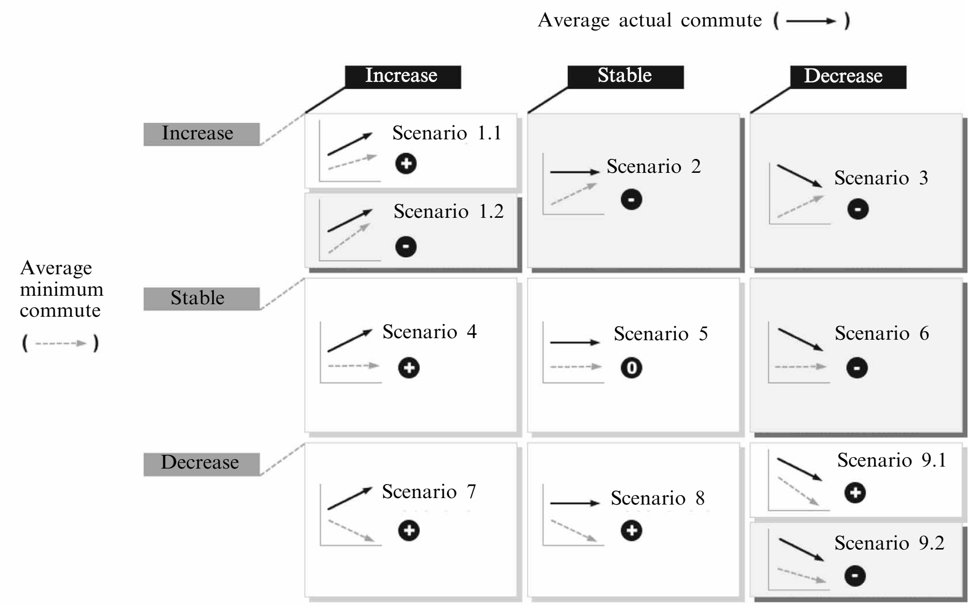
All the analyses discussed in this section are concerned with trip distance rather than time. This is because the influence of changing urban form on the estimation of the excess-commuting measure is better reflected when commuting distances are used as the cost measure. Time measure is not likely to be proportional to the distance (as speeds have increased), and so change over time would be lower, and this observation relates to the extensive literature on constant travel times (see Hupkes, 1982; Banister, 2005b). To highlight the structural changes in commuting patterns means that distance is the most appropriate metric to use, as evidenced from the work of Brotchie, Frost, and others. Hamilton (1989) has demonstrated that the use of a distance measure is related to upward bias, which produces larger amounts of excess commuting, whereas the use of travel times creates a downward bias. This is because the existence of fixed-time costs, such as parking time, is not considered in the transport-optimisation procedure, and longer journey distances are often related to the faster modes (for empirical evidence see Hamilton, 1989; Ma, 2004). Journey distance is the best measure of urban spatial change.

### 3.1 Eleven possible scenarios and a British case study

In this section we summarise the possible changes in the estimation of excess commuting that can be triggered by urban structural transformation and further investigate these changes in relation to a British case study by Frost et al (1998). A variety of combinations of changing average minimum commute and average actual commute can affect the results of excess commuting. Column 1 and row 1 (table 1) show the average minimum commute and average actual commute, respectively. It shows all the possible scenarios of the effects of changing urban form on the results of excess commuting when the approach is applied to one city over multiple time periods.

As indicated before, it is not always appropriate to argue that the reduced excess commute could be considered to be a good change (see the lightly shaded cells in the table). As shown in scenario 1.2 (table 1), an increase in minimum commute could cause a decrease in the amount of excess commute if an increase in actual commuting distance is slower than the minimum commute. (The opposite case is shown in scenario 9.2—a decrease in minimum commute may cause a reduction in the amount of excess commute if a decrease in actual commuting distance is faster than the minimum commute.) In relation to the previous discussion of monocentricity versus polycentricity, it means that monocentric development, which could cause an increase in the average

**Table 1.** Eleven scenarios of the effects of changing urban form on excess-commuting measures.




Notes:
(1) Changes of excess commuting are shown by + − 0 symbols and the table does not take into account the average maximum commute as excess commute is calculated from the difference between the theoretical minimum commutes and the actual commutes.
(2) The lightly shaded cells indicate a decrease in excess commuting with respect to urban-form changes.
(3) There are points where the proportion of excess travel remains stable somewhere in the middle between scenario 1.1 and 1.2, and between 9.1 and 9.2.
(4) Eleven possible scenarios are from nine possible combinations of average actual commute and average minimum commute. If the changes in maximum commute are taken into account, there are twenty-seven possible combinations. For simplicity, this table only presents the possible scenarios from nine possible combinations.

minimum commute, is likely to reduce the amount of excess commuting. The important point here is that, regardless of changes in average actual commute, such urban spatial changes could reduce the further potential of travel-length reduction by increasing the average minimum commute.

Scenario 1.2 seems to be one of the worst scenarios even though it shows a decrease in excess commuting. This scenario represents not only an increase in average actual commuting, but also a decrease in the possibility for workers to reduce their work-trip distances. It is worth noting that this occurs when the increase rate of the average minimum commute is greater than that of the average actual commute. Therefore, changes in the proportion of excess commuting should be taken into account with respect to the changes in minimum commute as well as those in actual commute.

From the transport planner's point of view, a decrease in actual commuting distance is one of the main objectives in achieving a sustainable transport goal. In this sense scenarios 3, 6, 9.1, and 9.2 seem desirable (table 1). However, once any changes of minimum values are considered, scenario 6 looks more desirable than scenario 3, and scenarios 9.1 and 9.2 are more desirable than scenario 6, as only these two scenarios provide the potential opportunity for workers to shorten their commuting distances. In particular, if the decrease in average actual commute is actual commute is equal, scenario 9.1 seems better than 9.2 even though it shows an increase in the amount of excess commuting. This is because scenario 9.1 could provide more potential opportunity to reduce the commuting length.

From these scenarios in table 1, it is clear that in a comparative study between different time periods, the change of the amount of excess commuting is not a good criterion for measuring urban-travel efficiency because negative or positive change in excess commuting does not contain any information about urban-travel efficiency. The empirical evidence of a British case study by Frost et al (1998), which is the only excess-commuting study where the analysis is applied at two points in time (between 1981 and 1991), provides a good example of explaining why a decrease in excess commuting does not always imply a movement towards efficient commuting travel. It seems that ten years is a sufficient time span to identify the changes of spatial distribution and their relation to excess commuting. Frost et al (1998) experimented with the excess-commuting method by including inward commuting. Inward commuting is undertaken by those with jobs in the city, but who live outside the city. The analysis tested the effects of the decentralisation of employees beyond the boundaries of each city over time. In addition, they estimated excess commuting under the conditions presented in most excess-commuting studies by excluding inward commuting, where only the work trips within the city boundaries are considered in the analysis.

The major findings of this study were that (1) the amount of excess commute is much smaller when inward commuting is incorporated into the transport optimisation model, and (2) the decentralisation of employment in the case-study cities had decreased the amount of excess commuting over the period between 1981 and 1991. Based on these findings, it was concluded that the result of the excess-commuting method is likely to be sensitive to the size of the study-area boundaries, especially when cities are compared within or between studies (Frost et al, 1998, page 537).

In table 2, the British case study by Frost et al (1998) shows a variety of possible scenarios when the traditional excess commuting method is applied to one city over multiple time periods. The number of scenarios identified is five out of eleven possible scenarios listed in table 1 (that is, scenarios 1.1, 1.2, 4, 7, and 9.1). When the inward commutes are included, all areas show an increase in both the average and minimum commutes during the period between 1981 and 1991. Except in the case of Birmingham, all case-study cities showed that the proportion of excess commuting

**Table 2.** Changes in commuting in a selection of British cities, 1981 – 1991 (source: Frost et al, 1998).

| | Jobs in 1981 | Average journey distance in 1981 (km) | Minimised journey distance in 1981 (km) | Excess commuting in 1981 (%) | Average journey distance in 1991[a] (km) | Minimised journey distance in 1991[a] (km) | Excess commuting in 1991[a] (%) | Change in share of inward commuting (%) | Change in average trip distance 1981 – 1991 (%) | Change in minimised trip distance 1981 – 1991 (%) | Change in proportion excess 1981 – 1991 (%) | Scenarios[b] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Including inward commuters* | | | | | | | | | | | | |
| London | 3 514 040 | 13.3 | 10.8 | 19.1 | 14.8 | 12.1 | 18.9 | 4.6 | 11.3 | 11.6 | −1.3 | 1.2 |
| Birmingham | 828 910 | 7.4 | 5.1 | 31.3 | 8.2 | 5.7 | 31.5 | 9.5 | 11.1 | 10.9 | 0.5 | 1.1 |
| Manchester | 831 260 | 7.4 | 5.0 | 32.0 | 8.7 | 6.1 | 29.8 | 17.5 | 17.3 | 21.0 | −6.9 | 1.2 |
| Liverpool | 459 070 | 7.3 | 5.0 | 32.0 | 7.9 | 5.5 | 31.2 | 7.5 | 8.3 | 9.5 | −2.6 | 1.2 |
| Newcastle | 310 830 | 7.7 | 5.8 | 24.5 | 8.6 | 6.6 | 22.5 | 7.7 | 11.2 | 13.5 | −8.1 | 1.2 |
| Leeds | 248 540 | 7.6 | 6.2 | 19.1 | 10.2 | 8.7 | 15.1 | 21.1 | 34.0 | 40.3 | −21.2 | 1.2 |
| Bristol | 249 520 | 7.5 | 5.7 | 24.4 | 9.3 | 7.5 | 19.1 | 14.8 | 23.6 | 32.1 | −21.7 | 1.2 |
| Nottingham | 199 870 | 6.6 | 5.1 | 22.3 | 7.6 | 6.5 | 13.8 | 15.1 | 15.6 | 28.3 | −38.2 | 1.2 |
| Sheffield | 229 400 | 6.4 | 4.8 | 25.9 | 7.5 | 6.1 | 20.5 | 17.7 | 17.9 | 26.5 | −20.8 | 1.2 |
| Leicester | 185 910 | 6.6 | 5.2 | 21.9 | 7.7 | 6.4 | 17.7 | 15.7 | 16.8 | 23.3 | −19.4 | 1.2 |
| *Excluding inward commuters* | | | | | | | | | | | | |
| London | 2 688 610 | 7.6 | 3.6 | 52.6 | 8.1 | 3.8 | 53.4 | - | 6.2 | 4.5 | 1.5 | 1.1 |
| Birmingham | 679 650 | 4.9 | 2.5 | 49.0 | 5.3 | 2.5 | 53.7 | - | 7.7 | −1.6 | 9.6 | 7 |
| Manchester | 718 760 | 4.8 | 2.3 | 52.1 | 5.1 | 2.2 | 54.7 | - | 6.9 | −4.3 | 5.0 | 7 |
| Liverpool | 395 370 | 4.6 | 2.2 | 55.2 | 4.8 | 2.2 | 56.1 | - | 5.4 | 2.2 | 1.7 | 1.1 |
| Newcastle | 229 600 | 4.0 | 2.0 | 50.0 | 4.3 | 2.0 | 54.6 | - | 8.2 | 0.0 | 9.2 | 4 |
| Leeds | 188 340 | 4.3 | 2.8 | 34.9 | 4.5 | 2.7 | 37.4 | - | 3.5 | −2.5 | 7.2 | 7 |
| Bristol | 196 650 | 3.7 | 1.9 | 48.6 | 3.8 | 1.9 | 50.5 | - | 3.2 | 0.5 | 3.9 | 1.1 |
| Nottingham | 151 680 | 3.6 | 1.9 | 47.2 | 3.7 | 1.9 | 49.5 | - | 3.0 | −2.1 | 4.9 | 7 |
| Sheffield | 173 920 | 3.9 | 2.1 | 46.2 | 4.0 | 2.1 | 45.9 | - | 2.3 | 2.3 | −0.6 | 1.2 |
| Leicester | 145 130 | 3.3 | 1.7 | 48.5 | 3.2 | 1.5 | 52.5 | - | −2.1 | −12.1 | 8.2 | 9.1 |

[a] These values were calculated, as they were not reported in the original paper.
[b] Refer to table 1 in the previous section.
Notes: Inward commuting is made by commuters who have jobs within the city but who reside outside. Frost et al's (1998) analysis was based on the assumption that the external zone where inward commuters originate from does not contain workplaces, therefore outward commuters are excluded.

declined because the minimum average commute increased faster than the actual commuting distance (note second column from right in table 2). All cities except Birmingham (scenario 1.1) conform to scenario 1.2 (table 1).

This finding is not surprising as the results derive mainly from the assumption that the external zone where inward commuters originate from does not contain workplaces. In fact, the assumption that there is an absence of suitable jobs in the city periphery is similar to assuming a monocentric urban-development pattern. It is easy to understand that under this monocentric-development assumption, further decentralisation of workers beyond the city boundaries tends to make both the minimised and actual travel distance greater than before. By taking into account the number of outward commuters, the amount of excess commuting is likely to increase through an increase in the average commute, and a decrease in the minimum commute.

However, as indicated earlier, despite a decrease in the proportion of excess commuting, such a monocentric development case (scenario 1.2 in table 1) does not seem to be a desirable change. First, they show an increase in the average commuting distance. Second, the relatively big increase of the minimum commuting distance means that spatial distribution of jobs and housing have developed in a potentially inefficient way, resulting in a reduction of the potential for work-travel savings.

In contrast, when the inward commutes are excluded (when journeys only within the city boundaries are taken into account), five different scenarios can be considered (scenarios 1.2, 1.2, 4, 7, 9.1). Five out of ten cities (Birmingham, Manchester, Leeds, Nottingham, and Leicester) show a decrease in the minimised travel distance. In relation to this, Frost et al (1998) noted "in some cities, the minimised travel distance shows a decrease over the 10 years implying the development of a potentially more efficient distribution of homes and jobs" (Frost et al, 1998, pages 536 – 537). However, among them, only Leicester, categorised into scenario 9.1, shows a decrease in the average actual commuting. Interestingly, in Leicester, there was an increase in the amount of excess commuting even though such a case indicates one of the best scenarios (table 1).

Four out of ten cities (Birmingham, Manchester, Leeds, and Nottingham) were categorised as scenario 7 in which cities show an increase in the average commuting distance despite a decrease in the minimum commuting distance. The situation here is that people travel further, even though the city form has become potentially more efficient. This conclusion is at odds with the general expectation that shorter commutes may be achieved by smaller minimum commutes associated with the mixed land use (for the discussion of the relationship between actual and minimum commutes see Giuliano and Small, 1993; Horner, 2002).

Probably, this interesting result is a good example of the situation where the urban flow is not determined solely by urban form. In the majority of excess-commuting studies, the estimation of the optimised commuting distance relies on the strong assumptions such as full knowledge, perfect rationality, flexible labour markets, single-worker household, and homogeneous housing type that facilitate the reallocation process. It is important to recognise that relaxing such strong assumptions (or imposing constraints) is related to a greater minimised commute, leading to more-conservative estimates of excess commute.

Obviously, certain factors are preventing workers from shortening their commuting and the issue of the efficient flow can go beyond the physical distribution of jobs and housing. Several major determinants such as multiworker household and the existence of the heterogeneity among workers were discussed in Hamilton (1982) and White (1988). Several empirical studies tried to take into account the possible constraints on the workers' mobility and produced less excess commuting (Buliung and Kanaroglou, 2002; Cropper and Gordon, 1991; Kim, 1994; Manning, 2003). More-complex analysis incorporating some of these additional constraints would help determine to what extent excess commuting contributes to the efficiency of urban travel.

**3.2 Simple simulation—the relationship between urban commuting capacity and travel**

It is clearly important that changes of excess commuting should be carefully considered in a comparative study between different time periods. Once again, this is simply because negative or positive change in excess commuting does not contain any information about the improvement in urban-travel efficiency. Investigating whether or not a particular area has transformed to a more efficient travel system needs to be discussed in relation to the urban-commuting potential determined by both the theoretical minimum and maximum commutes. As explained in section 2.2, urban-commuting potential is comprised of further potential increases and decreases. Therefore, the possibility of how much an actual trip may increase (or decrease) can be discussed in the framework of the notion of urban-commuting potential.

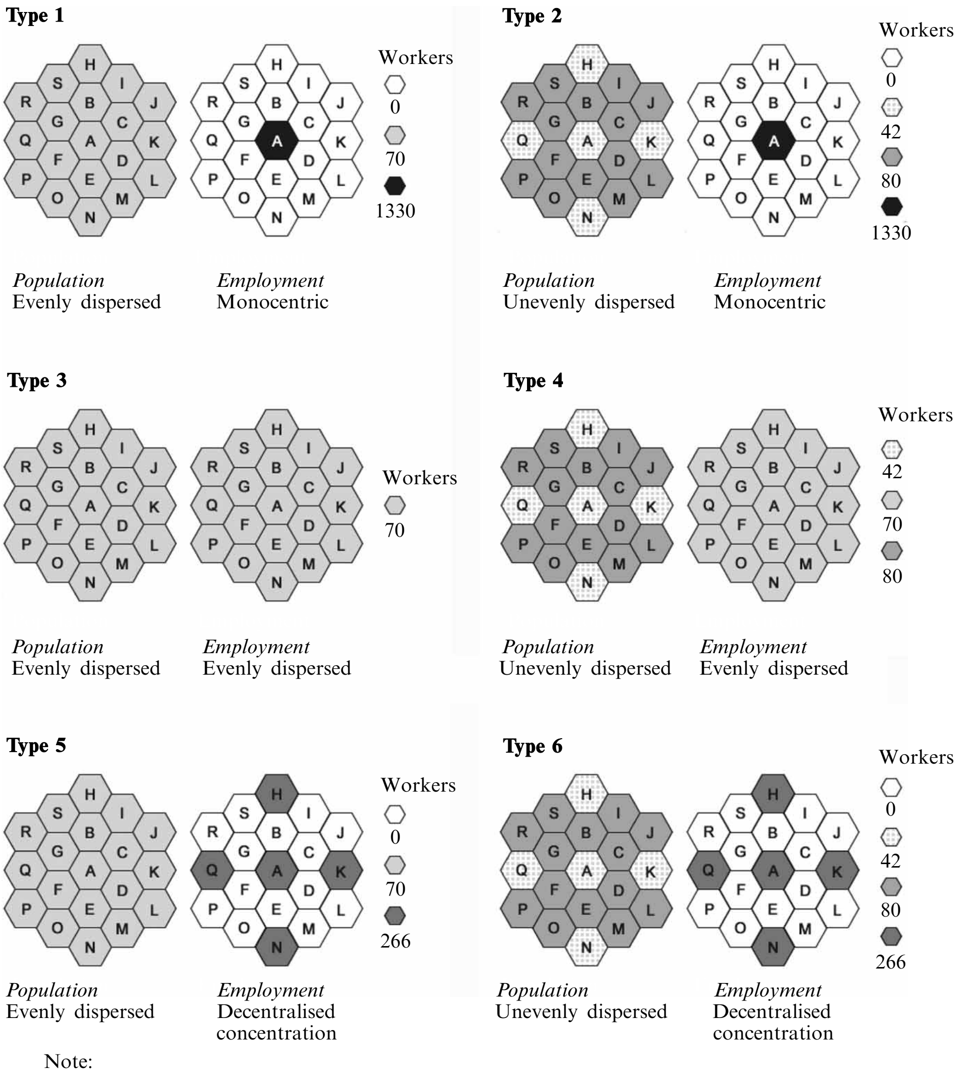
Horner (2002) asserted that the city with a small difference between maximum and actual commutes is considered to be an inefficient one. In other words, an efficient work-travel situation is reflected by a large difference between maximum and actual commutes in the urban-commuting potential (or by a small difference between minimum and actual commutes). Indeed, the travel efficiency could be discussed in relation to how much people optimise their travel in a given travel capacity, as the concept of the term 'efficiency' means the ratio of the effective result without wasted time or effort in a total-system capacity.

In this sense, the concept of the commuting potential is very important, and in this section we now further investigate the relationship between changing urban form and the theoretical maximum and minimum values in the extended excess-commute measures. Two important points are emphasised in this section in relation to the difference between the standard and the extended excess-commuting method: (1) the commuting potential varies according to different urban spatial structures and (2) urban decentralisation may play a significant role in changing urban-commuting potential. These are the important aspects that the extended excess-commuting measure can capture.

Six simulations are conducted to examine their effect. Figure 4 shows six different types of urban patterns. Each city type has the possibility of changing from one type to another over time. In this figure, type 1 and type 2 present the simplest monocentric urban model, in which all workers are in the central business district but have different population distributions. Evenly dispersed employment distributions are shown in type 3 and type 4. In addition, type 5 and type 6 are introduced to present the decentralised dispersion of employment.

Urban space in this simulation consists of nineteen zones that are represented by nineteen hexagonal shapes. In addition, the simulation uses the assumption of universality in the transport system. Some of the assumptions used here are noted at the bottom of figure 4. Under these assumptions, the theoretical maximum commute (upper bound) and minimum commute (lower bound) for each city pattern are calculated, using the extended excess-commuting measure (table 3).

The amount of excess commuting is zero under complete employment centralisation. In the monocentric city, the minimum commutes are the same as the maximum and actual commutes because it is impossible to trade jobs and residential places in the optimisation process. This reconfirms our previous conjecture that a city with a relatively high concentration of population or employment in a central area is likely to have smaller amounts of excess commute regardless of actual commuting distance. In a completely monocentric structure, the only way to reduce the actual commuting




Note:
The assumptions used in this simulation are as follows:
(1) Urban space is made up of nineteen zones and every zone is the same size,
(2) the commuting distance between the neighbouring zones is 10 km,
(3) the commuting distance within each zone is 5 km,
(4) for convenience of calculation, the maximum distance from the centre to the boundary areas is 30 km; therefore, 19 zones are used in the simulation,
(5) all workers have identical jobs,
(6) for convenience of calculation, it is assumed that the total number of workers is 1330 (note that the number of workers used here does not affect the result—average commuting distances),
(7) five urban centres were assumed in the city form of decentralised concentration. In an unevenly dispersed population model, the size of the population in some parts was assumed to be approximately double that of the other parts.

**Figure 4.** Six different types of city forms.

distance (or the theoretical maximum and minimum commutes) is to move workers closer to their workplaces (from type 1 to type 2).

In contrast, decentralisation, or further dispersion of city form, shows an influence on the excess commuting measures. Further transformation from a monocentric to a polycentric city makes a lower bound on the excess-commute measure lower, and an

**Table 3.** Maximum and minimum commutes in different city forms.

| Employment distribution | Minimum ~ maximum commute (km) | |
|---|---|---|
| | evenly distributed population | unevenly distributed population |
| Monocentric | 16.05 (type 1) | 15.92 (type 2) |
| Evenly dispersed | 5.00 ~ 31.85 (type 3) | 5.53 ~ 31.71 (type 4) |
| Decentralised concentration | 8.68 ~ 32.05 (type 5) | 9.21 ~ 31.92 (type 6) |

Notes:
(1) The theoretical maximum commute and minimum commute are based on the separate calculation (see appendix for the computation of these bounds).
(2) Type 3 has the largest commuting potential range (max − min = 26.85), followed by type 4 (max − min = 26.18).

upper bound higher (from type 1 to type 5, and from type 2 to type 6). These results indicate that, in general, the higher degree of decentralisation in a city is likely to produce much larger excess-commuting values in the transport-optimisation modelling because it diminishes the theoretical minimum commutes (note that the amount of excess commuting is zero in types 1 and 2). However, it seems that an evenly dispersed city structure does not always produce the highest value of the theoretical maximum commute. A decentralised concentration of city form could either have higher or lower values than an evenly dispersed city. This varies according to the constraints, or the actual distribution of jobs and residential locations.

Now, to compare the excess-commuting values between different city types, let us suppose that the average actual commuting distances are 16.05 km in type 1, type 3, and type 5, and 15.92 km in type 2, type 4, and type 6 (the average actual commuting distances used here are based on the monocentric employment distribution—type 1 and type 2 in table 3). In this example, the theoretical amounts of excess commutes of type 1 to type 6 are 0%, 0%, 69%, 65%, 46%, and 42%, respectively (calculated by the difference between the maximum and minimum divided by the average). It is very interesting to note that even though six types of models in figure 4 have the same size and number of workers, and similar actual commuting distances, there is much variation in the differences of the extended excess-commuting values between models.

From these simple simulations, it should be emphasised that the cities with a high proportion of excess commuting also have a high possibility of reducing average actual commuting. For the comparative analysis over multiple time periods, it is desirable that the change of average actual commute is considered with the change in the extended excess commuting and the commuting potential.

## 4 Conclusions

The estimation of excess commuting has clearly shown that the conventional excess-commuting measurement alone might not be appropriate to describe urban-travel efficiencies when it is used in the same city over different time periods. A further investigation of this finding is that careful interpretation is required when the excess-commuting values are compared between different cities, as cities have different urban structures and commuting potential. It should be emphasised that smaller excess-commuting values do not necessarily mean better commuting efficiency.

The introduction of the maximum commute in the conventional excess commuting method allows the urban commuting potential to be investigated in relation to urban form. When the extended excess commuting method is applied to different time periods, this can be used as a useful application tool for benchmarking commuting efficiency of a particular city. Indeed, capturing the relationship between the dynamics of population and employment in an urban area and commuting efficiency is one of the most novel aspects of the extended excess-commuting method described in this paper.

Two other points need to be made. In this paper the focus has been on the commuting journey, as this is a major trip-making activity and is often seen as a key determinant in where people choose to live and work. But the commuting journey is declining in importance in terms of its share of all trip activity. For example, in 1985/86 [Frost et al's (1998) comparative study period] the journey-to-work trip only accounted for 17.4% of trips and 20.4% of trip distance in the United Kingdom. The corresponding figures for 1999/2001 for the journey to work are 15.3% of trips and 19.4% of distance (Department for Transport, 2004). Other trip purposes such as leisure, social, and shopping have become more important. This means that reducing the gap between the actual average and minimum commutes should not be a policy target on its own, as this may decrease overall urban-travel efficiencies by increasing other-purpose trip lengths. A study of the purposes of other trips would provide evidence on this proposition, provided that suitable data are available.

The second point relates to new patterns of working, with the acceptance of the heterogeneity of the job market, the importance of schools, and other factors in location decisions, and the increase in multiple workers in households, with the increase in part-time and female participation in the labour force. All these factors create a greater complexity in the analysis, but their inclusion might lead to a greater realism in the analysis (provided that the data are available). The policy implications here are that mixed land uses and concentration will not necessarily on their own result in reductions in excess commuting. The key socioeconomic variables also need to be included to cover the full range of explanatory variables when examining overall levels of excess travel (including excess commuting). The degree of mixture between jobs and housing only represents the potential for shorter (or longer) journeys in the process of decentralisation, and the evidence provided by Crane and Chatman (2003) clearly supports this idea that the effects of decentralisation on commutes may be better understood, when they are considered together with other demographic and economic characteristics of the individuals and households.

Urban decentralisation can lead to an increase or a decrease in average commuting distance and this challenge has been an intriguing issue for many researchers. As Crane and Chatman (2003) noted, however, real evidence about the relationship between decentralisation and commuting distance is surprisingly rare. Further research is needed to explore how the dynamics of excess commuting can be matched with changes in actual commuting behaviour for several cities with different structures and development patterns, and how other trip purposes can be included so that a more complete picture of change can be built up.

**References**

Anas A, Arnott R, Small K A, 1998, "Urban spatial structure" *Journal of Economic Literature* **36** 1426 – 1464

Anderson W P, Kanaroglou P S, Miller E J, 1996, "Urban form, energy and the environment: a review of issues, evidence and policy" *Urban Studies* **33** 7 – 35

Banister D, 2005a *Unsustainable Transport: City Transport in the New Century* (Routledge, London)

Banister D, 2005b, "Time and travel", in *Methods and Models in Transport and Communications: Cross Atlantic Perspectives* Eds A Reggiani, L Schintler (Springer, Berlin)
Baumont C, Ertur C, Le Gallo J, 2003, "Spatial analysis of employment and population density: the case of the agglomeration of Dijon, 1999", http://econwpa.wustl.edu/
Bertaud A, 2002, 'Note on transportation and urban spatial structure", paper presented at the Annual Bank Conference on Development Economics, Washington, DC, copy available from http://alain-bertaud.com
Bourne L S, 1982, "Urban spatial structure: an introductory essay on concepts and criteria", in *Internal Structure of the City* Eds L S Bourne (Oxford University Press, New York) pp 28 – 45
Brotchie J F, 1984, "Technological change and urban form" *Environment and Planning A* **16** 583 – 596
Brotchie J F, Anderson M, Gipps P G, McNamara C, 1996, "Urban productivity and sustainability—impact of technological change", in *Transport, Land-use and the Environment* Eds Y Hayashi, J Roy (Kluwer, Dordrecht)
Buliung R N, Kanaroglou P S, 2002, "Commute minimization in the Greater Toronto Area: applying a modified excess commute" *Journal of Transport Geography* **10** 177 – 186
Crane R, 2000, "The influence of urban form on travel: an interpretive review" *Journal of Planning Literature* **15** 3 – 23
Crane R, Chatman D G, 2003, "Traffic and sprawl: evidence from US commuting 1985 to 1997" *Planning and Markets* **6** 14 – 22
Cropper M, Gordon P, 1991, "Wasteful commuting: a re-examination" *Journal of Urban Economics* **29** 2 – 13
Department for Transport, 2004 *National Travel Survey* (The Stationery Office, London)
Frost M, Linneker B, Spence N, 1998, "Excess or wasteful commuting in a selection of British cities" *Transportation Research A* **32** 529 – 538
Giuliano G, Small A K, 1993, "Is the journey to work explained by urban structure?" *Urban Studies* **30** 1485 – 1500
Gordon P, Wong H L, 1985, "The costs of urban sprawl: some new evidence" *Environment and Planning A* **17** 661 – 666
Hamilton B W, 1982, "Wasteful commuting" *The Journal of Political Economy* **90** 1035 – 1053
Hamilton B W, 1989, "Wasteful commuting again" *Journal of Political Economy* **97** 1497 – 1504
Horner M W, 2002, "Extensions to the concept of excess commuting" *Environment and Planning A* **34** 543 – 566
Hupkes G, 1982, "The law of constant travel time" *Futures* **14** 38 – 46
Kim S, 1995, "Excess commuting for two-worker households in the Los Angeles metropolitan areas" *Journal of Urban Economics* **38** 166 – 182
Klaassen L H, Paelinck J H P, 1981 *Dynamics of Urban Development* (Gower, Aldershot, Hants)
Ma K, 2002, "The changes of urban spatial structure and commuting in Seoul", paper presented at the Universities' Transport Study Group 34th Annual Conference, Edinburgh; copy available from the Transport Research Institute, Napier University, Edinburgh
Ma K, 2004 *The Impact of Urban Spatial Decentralisation on Jobs-Housing Imbalance* unpublished PhD thesis, The Bartlett School of Planning, University College London, London
Manning A, 2003, 'The real thin theory: manopsony in modern labour markets" *Labour Economics* **10** 105 – 131
Merriman D, Ohkawara T, Suzuki T, 1995, "Excess commuting in the Tokyo metropolitan area: measurement and policy simulations" *Urban Studies* **32** 69 – 85
Newman P W F, Kenworthy J R, 1992, "Is there a role for physical planners?" *Journal of American Planning Association* **58** 353 – 362
O'Sullivan A M, 1999 *Urban Economics* (McGraw-Hill, London)
Scott D, Kanaroglou P, Anderson W, 1997, "Impact of commuting efficiency on congestion and emissions: case of the Hamilton CMA, Canada" *Transportation Research D* **2** 245 – 257
White M J, 1988, "Urban commuting journeys are not 'wasteful'" *Journal of Political Economy* **96** 1097 – 1110

## Appendix

The computation of the average actual and the theoretical minimum commuting distances is known as a transport problem. The actual total commuting can be given by

$$\sum_i \sum_j c_{ij} n_{ij} ,$$

where $c_{ij}$ is the commuting distance between zone $i$ and zone $j$ and the actual number of workers who live in zone $i$ and commute to zone $j$ is given $n_{ij}$. The optimised journey-to-work matrix containing $n_{ij}^*$ can be obtained by choosing the number of workers who should live in $i$ and work in $j$ in the linear-programming procedure. The minimum total commute can be calculated by

$$\sum_i \sum_j c_{ij} n_{ij}^* .$$

On the other hand, the maximised journey to work matrix containing $n_{ij}^*$ can be obtained by

$$\max \sum_i \sum_j c_{ij} n_{ij} ,$$

which can also be expressed by

$$\left| \min \left( \sum_i \sum_j -c_{ij} \cdot n_{ij} \right) \right|$$

according to the simplex method in linear programming (see Horner, 2002).